

# Notice of Allowability

Application No.

10/675,756

Examiner

DIANE D. MIZRAHI

Applicant(s)

NEVILL-MANNING ET AL.

Art Unit

2165

## -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to 12-5-06.
2. ☒ The allowed claim(s) is/are 3-12, 15, and 19-24 (renumbered 1-17).
3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a) ☐ All b) ☐ Some\* c) ☐ None of the:
    1. ☐ Certified copies of the priority documents have been received.
    2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

\* Certified copies not received: \_\_\_\_\_.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.

**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
  - (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
    - 1) ☐ hereto or 2) ☐ to Paper No./Mail Date \_\_\_\_\_.
  - (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date \_\_\_\_\_.Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

### Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date \_\_\_\_\_
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☒ Interview Summary (PTO-413), Paper No./Mail Date 12-19-06.
7. ☐ Examiner's Amendment/Comment
8. ☐ Examiner's Statement of Reasons for Allowance
9. ☐ Other \_\_\_\_\_.

DIANE D. MIZRAHI  
PRIMARY EXAMINER

### EXAMINER'S AMENDMENT

An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Attorney Brenda Simon on

December 19, 2006.

*The application has been amended as follows:*

1. (Canceled).
2. (Canceled).
3. (Currently Amended) The method of claim [[2]] 12, wherein comparing the seed article to at least one other related article is performed by a dynamic programming alignment algorithm to determine an alignment between the seed article and the related article.
4. (Currently Amended) The method of claim [[2]] 12, further comprising determining a cluster of related articles from the related articles.
5. (Previously Presented) The method of claim 4, wherein determining the cluster of related articles is performed by:

using a dynamic programming alignment algorithm to compute edit distances between the seed article and the related articles; and  
choosing the cluster of related articles based on the edit distances.

6. (Original) The method of claim 4, wherein the identifying at least one information field within the seed article is performed by comparing the seed article to the cluster of articles.

7. (Currently Amended) The method of claim [[2]] 12, wherein the information field corresponds to variable data.

8. (Currently Amended) The method of claim [[2]] 12, wherein the articles are web pages.

9. (Original) The method of claim 8, wherein the related articles are web pages on a web site.

10. (Original) The method of claim 9, further comprising simplifying the content on a web page.

11. (Original) The method of claim 10, wherein simplifying the content includes preserving visible text, visible images, and visible paragraph and table formatting.

12. (Currently Amended) ~~The method of claim 2, further comprising:~~ A method for information extraction, comprising:

accessing a plurality of related articles;

determining a seed article from the related articles;

identifying at least one information field within the seed article by comparing the seed article to at least one other related article;

creating a template based on the identified information field;

identifying a plurality of templates each comprising at least one information field;

comparing a source article to the templates to determine a closest template;

associating data from the source article with an information field from the closest

template; and

extracting the associated data.

13. (Canceled).

14. (Canceled).

15. (Currently Amended) A method of extracting data from a source article,

comprising:

identifying a plurality of templates each comprising at least one information field;

comparing the source article to the templates to determine a closest template, wherein

comparing the source article to the templates is performed by a dynamic

programming alignment algorithm to compute an edit distance between the source

article and the templates;

associating data from the source article with an information field from the closest

template;

extracting the associated data; and

displaying the associated data.

16. (Canceled).

17. (Canceled).

18. (Canceled).

19. (Currently Amended) The computer program product of claim ~~[[18]]~~23, wherein comparing the seed article to at least one other related article is performed by a dynamic programming alignment algorithm to determine an alignment between the seed article and the related article.

20. (Currently Amended) The computer program product of claim ~~[[18]]~~23, further comprising computer program code for determining a cluster of related articles from the related articles.

21. (Previously Presented) The computer program product of claim 20, wherein determining a cluster of related articles is performed by:  
using a dynamic programming alignment algorithm to compute edit distances between  
the seed article and the related articles; and  
choosing the cluster of related articles based on the edit distances.

22. (Previously Presented) The computer program product of claim 20, wherein the identifying at least one information field within the seed article is performed by comparing the seed article to the cluster of related articles.

23. (Currently Amended) ~~The computer program product of claim 18, further comprising computer program code for:~~ A computer program product for information extraction, comprising:

a computer-readable medium; and

computer program code, encoded on the medium, for:

accessing a plurality of related articles;

determining a seed article from the related articles;

identifying at least one information field within the seed article by comparing the seed article to at least one other related article;

creating a template based on the identified information field;

identifying a plurality of templates each comprising at least one information field;

comparing a source article to the templates to determine a closest template;

associating data from the source article with an information field from the closest template; and

extracting the associated data.

24. (Previously Presented) The computer program product of claim 23, wherein comparing the source article to the templates is performed by a dynamic programming alignment algorithm to compute an edit distance between the source article and the templates.

**Allowable Subject Matter**

Claims 3-12, 15, and 19-24 are allowed over the prior art made of record.

**Comments**

The date on which the petition under 37 CFR 1.136(a) and the appropriate extension fee have been filed is the date for purposes of determining the period of extension and the corresponding amount of the fee. In no case may an applicant reply outside the SIX (6) MONTH statutory period or obtain an extension for more than FIVE (5) MONTHS beyond the date for reply set forth in an Office action. A fully responsive reply must be timely filed to avoid abandonment of this application.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue

fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

As allowable subject matter has been indicated, Applicant's response must either comply with all formal requirements or specifically traverse each requirement not complied with. See 37 CFR 1.111(b) and MPEP section 707.07(a).

**Other Prior Art Made of Record**

The prior art made of record and not relied upon is considered pertinent to Applicant's disclosure. U.S. patents and U.S. patent application publications will not be supplied with Office actions. Examiners advises the Applicant that the cited U.S. patents and patent application publications are available for download via the Office's PAIR. As an alternate source, all U.S. patents and patent application publications are available on the USPTO web site ([www.uspto.gov](http://www.uspto.gov)), from the Office of Public Records and from commercial sources. For the use of the Office's PAIR system, Applicants may refer to the Electronic Business Center (EBC) at <http://www.uspto.gov/ebc/index.html> or 1-866-217-9197.

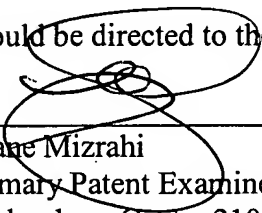
**Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Diane D. Mizrahi whose telephone number is 571-272-4079. The examiner can normally be reached on Monday-Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey Gaffin can be reached on (571) 272-4146. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 872-9306 for regular communications and (703) 305-3900 for After Final communications.

Art Unit: 2165

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-3900.



---

Diane Mizrahi  
Primary Patent Examiner  
Technology Center 2100

December 19, 2006